IAFF 6501- 11

Applied Quantitative Analysis

Welcome!

- Please pick up your name tent
- The Midterm Exam Review Guide has been posted to course blackboard

Class 5
- ***Review of contingency tables***
- ***More on z scores***
- ***Sampling Distribution***
- ***Standard Error***
- ***Confidence Intervals***
- ***t distribution***

Handedness of Male and Females

Cell counts

	Female	Male	total
Right	8	7	15
Left	3	1	4
total	11	8	19

Row percents: Handedness by Gender

	Female	Male	total
Right	53%	47%	100%
Left	75%	25%	100%

Column percents: Gender by Handedness

	Female	Male	
Right	73%	88%	
Left	27%	13%	
total	100%	100%	

SPSS Example

2 Variables: SEX and CASINO

	Female	Male	total
Agree			
Disagree			

Casino by Sex

	Female	Male	total
Agree			100%
Disagree			100%

Row %s

Sex by Casino

	Female	Male	total
Agree			
Disagree			
	100%	100%	

Column %s

Assignment #1

2 Variables: SEX and SPIDERS

	Female	Male	total
Agree			
Disagree			

SPIDERS by Sex

	Female	Male	total
Agree			
Disagree			

Sex by SPIDERS

	Female	Male	total
Agree			
Disagree			

Recap from last week

Probability distributions

The standard normal curve provides us with a scale for the likelihood of being above or below any point on the scale. The *z*-score indicates position in a normal distribution in terms of the number of standard deviations that a score is above or below the mean.

Z score is known as the "standard score"

Given a population mean (μ) and population standard deviation (σ), we can calculate z-score for any observed score (x), and vice versa.

$$z = (x - \mu) / \sigma$$

$$x = \mu + z\sigma$$

We can use normal curve probability table to interpret a z score in terms of probabilities (probability of finding a score above or below that value).

Practice with z-scores

1. If value x has a z= + 1.34, what percent of values will fall above x? How do you interpret this?
   - 9 percent of values will fall above that x
   - .09 probability of getting a value this high or higher

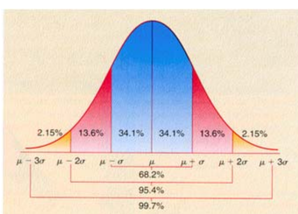
1. If z= - 0.77, what percent of values will fall below x? How do you interpret this?
   - 22 percent of cases will fall below that x
   - .22 probability of getting a value this low or lower

More Practice with z-scores

Practice problem: The average home size in a region is 900 square feet. Assume that home size is normally distributed in this region, and that the standard deviation of this distribution is 120 square feet. About what percent of the homes in this region are smaller than 740 square feet?

Standard Normal Distribution

- The normal distribution with $\mu = 0$, $\sigma = 1$
- For that distribution, number of falling z standard deviations above the mean is $\mu + z\sigma = 0 + z(1) = z$; its simply the z-score itself
  - z = 2 is 2 standard deviations above the mean
  - z= 1.3 is 1.3 standard deviations above the mean.



Standard Normal Distribution

Why is normal distribution so important?

We' ll learn today that if different studies take random samples and calculate a statistic (e.g. sample mean) to estimate a parameter (e.g. population mean), the collection of statistic values from those studies usually has approximately a normal distribution.

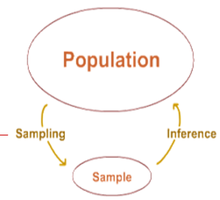
9

Sample, Population, & Sampling

Sampling Procedure

Random – every member in a population has an equal chance of being drawn into the sample
- Simple Random
- Multistage

Nonrandom – not every member in a population has a equal chance of selection
- Accidental
- Quota
- Judgment or Purposive



10

Sampling Distribution

- lists the possible values of a statistic (e.g., sample mean or sample proportion) and their probabilities
- Suppose we have a population: 6 7 8 9 ; $\mu = 7.5$
- For possible samples of size $n = 2$, consider the sample mean

1.	6	6	$\bar{X}=6.0$	7.	7	8	$\bar{X}=$ **7.5**	13.	9	6	$\bar{X}=$ **7.5**
2.	6	7	$\bar{X}=6.5$	8.	7	9	$\bar{X}=8.0$	14.	9	7	$\bar{X}=8.0$
3.	6	8	$\bar{X}=7.0$	9.	8	6	$\bar{X}=7.0$	15.	9	8	$\bar{X}=8.5$
4.	6	9	$\bar{X}=$ **7.5**	10.	8	7	$\bar{X}=$ **7.5**	16.	9	9	$\bar{X}=9.0$
5.	7	6	$\bar{X}=6.5$	11.	8	8	$\bar{X}=8.0$				
6.	7	7	$\bar{X}=7.0$	12.	8	9	$\bar{X}=8.5$				

11

Sampling Distribution of the Mean

- Definition: The probability distribution of the sample means derived from all possible random samples of a given size from a given population.
- The sampling distribution of the mean informs us of the degree of sample-to-sample variability we should expect due to chance.

12

Importance of Sampling Distributions

- Tell us the probability of getting a particular sample mean, given μ & σ
- Critical for inferential statistics
- Allow us to …
  - Estimate population parameters (μ & σ)
  - Determine if a sample mean differs from a known population mean just because of chance
  - Compare differences between sample means due to chance

13

Sampling Distribution of the Mean
Important Characteristics

1. Mean approximates a normal curve
2. Mean of the sampling distribution of means is equal to the true population mean
3. The standard deviation of sampling mean less than standard deviation of the population

14

Sampling Distribution of the Mean & Normal Curve

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

where: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

15

Practice Problem

Suppose you take a random sample of 25 graduate students, and measure their IQ. Assuming that IQ is normally distributed, you get a sample mean equal to 105. However, you are not sure if that is accurate, and want to know the probability of the sample mean being greater than 105

$\mu = 100$ and $\sigma = 15$.

16

Practice Problem Answer

Step 1: Convert to Z-score $z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$

$$\sigma_{\bar{x}} = \frac{15}{\sqrt{25}} = \frac{15}{5} = 3$$

$$z = \frac{105 - 100}{3} = 1.67$$

Step 2: find our z-score of 1.67 in table A in appendix C

we see that the probability of the sample having a mean of 105 or greater is: .0475 *(I converted the 4.75% into a proportion).*

17

Standard Error of the Mean

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Example:

- *Mean = 169*
- *$\sigma = 5$*
- *$n = 30$*
- *$\sigma_x = .91$*

- What does this mean about our precision/accuracy/reliability in our estimates?

Confidence Intervals (CI)

$$CI_\mu = \bar{X} \pm \underbrace{Z\sigma_{\bar{X}}}_{\text{Margin of Error}}$$

Confidence Intervals and Z-scores

- 68% CI has Z = $\pm$ 1.00
- 95% CI has Z = $\pm$ 1.96
- 99% CI has Z = $\pm$ 2.58

Example

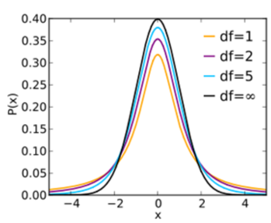
Let's assume they we have a sample of 100 and we find sample mean equal to 26 and σ= 4 and we want to create a 95% confidence interval for the population mean.

Step 1: Calculate the Standard Error of the Mean 4/ √100 = .4

Step 2: Calculate the Margin of Error 1.96 (.4) = .78

Step 3: Add and subtract the margin of error from the sample mean 26+- .78= 25.22 to 26.78

t distribution



20

t distributions

- allows us to conduct statistical analyses on data sets that are not appropriate for using the normal distribution

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{N\text{-}1}}}$$

Standard error of the mean

21

s and s^2 Unbiased Formulas

Sample Variance	Sample Standard Deviation
$s^2 = \frac{\sum(x-\bar{x})^2}{n-1}$	$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Use these formulas!!!

22

Degrees of Freedom

- Definition - the number of independent observations in a set of data
- $df = N - 1$
- Table C in appendix
- As sample size increases, *df* increases and t distribution gets closer to the normal distribution

23

Northeastern University
Online

Table C

- Calibrated for various levels of α
- α = 1 – level of confidence
  95% level of confidence α = .05
  99% level of confidence α = .01
- Example Using Table C:
  - N = 20, df = 29, α =.05, t = 2.045
- Practice : *N* = 20, α =.01, what is *t* value?

24

Questions

25

Midterm Exam ~ Oct. 10th

- 1 page (8.5 by 11) of handwritten notes allowed
  - Formula
  - Statistical symbols
- Appendix Tables will be provided
- Review guide posted to BB
- Review practice problems
- 1 hour exam